Transcription of *Rhodospirillum rubrum* *atp* operon

Gunnar FALK* and John E. WALKER†

M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

(Received 10 January 1985/15 March 1985; accepted 4 April 1985)

The photosynthetic non-sulphur bacterium *Rhodospirillum rubrum* contains a cluster of five genes encoding the subunits of F_1 -ATPase [Falk, Hampe & Walker (1985) *Biochem. J.* **228**, 391–407]. Transcription of these genes has been studied by two methods, transcriptional mapping with S_1 nuclease and primer extension analysis. Thereby a 5'-end in RNA derived from this region has been demonstrated at a guanine residue 236 bases before the initiation codon of the gene for the δ -subunit, the first in this cluster. DNA sequences on the 5' side of this nucleotide show some similarity to promoters in *Escherichia coli*, but are not apparently related to sequences upstream of the *Rhodopseudomonas blastica* *atp* operon. A 3'-end in RNA derived from this gene cluster has been demonstrated by S_1 -nuclease mapping. This is found before a run of thymidylate residues in the DNA, on the 3' side of a region of dyad symmetry. In *E. coli* these features are characteristic of *rho*-independent transcriptional termination signals. It appears from these studies and from the organization of the genes that the five genes in the *atp* cluster may be co-transcribed from this promoter and that transcripts terminate at the region of dyad symmetry.

The five genes encoding the subunits of *Rhodospirillum rubrum* F_1 -ATPase form a cluster, as demonstrated by DNA sequence analysis (Falk *et al.*, 1985). A similar arrangement is found in the related micro-organism *Rhodopseudomonas blastica* (Tybulewicz *et al.*, 1984), and a cluster of genes for F_1 -subunits form part of the *Escherichia coli* *unc* operon (Walker *et al.*, 1984a) (see Fig. 1). Transcription of the nine genes of the *unc* operon proceeds from a single promoter. The transcriptional initiation site is precisely known (Porter *et al.*, 1983; Walker *et al.*, 1984a), and sequences upstream of this site are related to other *E. coli* promoters (Gay & Walker, 1981). Transcription of the *Rps. blastica* *atp* operon also proceeds predominantly from a single site, which is upstream of the gene for the δ -subunit. Some transcription also arises from a second site between the γ - and β -genes (Tybulewicz *et al.*, 1984). However, sequences on the 5'-side of these transcriptional start sites are not apparently related to *E. coli* promoter sequences. This observation is consistent with the finding that DNA templates from a

related photosynthetic bacterium, *Rhodopseudomonas sphaeroides*, are not transcribed and translated *in vitro* by extracts of *E. coli*, whereas *E. coli* DNA templates are (Chory & Kaplan, 1982). Beyond the gene for the ϵ -subunit, which is promoter distal in the *E. coli* and *Rps. blastica* operons, are sequences with dyad symmetry that could form stable hairpin-loop structures, followed by runs of thymidylate residues. These features are characteristic of *rho*-independent transcriptional termination signals (Pribnow, 1978), and transcription of the *Rps. blastica* *atp* operon has been shown experimentally to terminate at this site (Tybulewicz *et al.*, 1984).

In the present paper we describe experiments that define the sites of initiation and termination of transcription of the *Rsp. rubrum* *atp* gene cluster. These experiments suggest that the five genes are co-transcribed from a single promoter. Transcription terminates at a sequence that resembles an *E. coli* *rho*-independent transcriptional terminator.

Materials and methods

Materials

Sources of materials, chemicals and most biochemicals have been described (Falk *et al.*, 1985). S_1 nuclease was purchased from Bethesda Re-

* Present address: Department of Biochemistry, Arrhenius Laboratory, University of Stockholm, Sweden.

† To whom correspondence should be addressed.

Fig. 1. Arrangement of genes for prokaryotic H^+ -translocating ATPases

The genes are named according to the subunits they encode, except for the *E. coli* *uncI* gene (Gay & Walker, 1981; Gay, 1984) and the *Rps. blastica* X gene (Tybulewicz *et al.*, 1984) both of which encode proteins of unknown function. The scale is in kilobases.

search Biochemicals (Bethesda, MD, U.S.A.) and reverse transcriptase from Anglian Biotechnology (Colchester, Essex, U.K.).

Preparation of RNA

Rsp. rubrum (strain S1) was grown aerobically for 40 h at 30°C in 50 ml of the following rich medium in a 250 ml conical flask: 0.3% yeast extract/0.3% peptone/2 mM- $CaCl_2$ /2 mM- $MgSO_4$, pH 6.8. The organism was also grown anaerobically at room temperature for 48 h in the same medium in a screw-cap bottle (10 ml) with illumination with white light. The cultures were centrifuged, and the pellets were resuspended in 4 ml of lysis buffer (0.5% sodium dodecyl sulphate/20 mM-sodium acetate/1 mM-EDTA, pH 5.0). The resulting solutions were extracted twice at 60°C with phenol (5 ml) that had been equilibrated in lysis buffer. Nucleic acid was precipitated and washed with ethanol, dried *in vacuo*, redissolved in water (500 μ l) and stored at -20°C.

Mapping with S_1 nuclease

Single-stranded probes for mapping the mRNA derived from the *atp* region were made by the prime-cut method (Farrell *et al.*, 1983), as described previously (Tybulewicz *et al.*, 1984). The probe (5×10^4 c.p.m.) and the bacterial RNA solution (6 μ l) were co-precipitated with ethanol and then resuspended in the annealing buffer containing 80% (v/v) formamide (Berk & Sharp, 1978). The RNA was denatured at 80°C for 10 min and then hybridized with the probe at 57°C for 12–16 h. Then complexes were digested with S_1 nuclease by the addition of 200 μ l of an ice-cold solution containing buffer [30 mM-sodium acetate (pH 4.5)/0.25 M-NaCl/1 mM- $ZnSO_4$ /5% (v/v) glycerol] and S_1 nuclease (400 units). The solution was incubated at 37°C for 30 min. Then the DNA was precipitated with ethanol and redissolved in formamide (4 μ l). The sample was boiled for 3 min and then analysed by electrophoresis on a 6% (w/v) polyacrylamide gel.

Primer extension analysis

Prime-cut probes made from appropriate clones in bacteriophage M13 (see Table 1) were annealed to RNA in 80% formamide at 57°C as described above for S_1 -nuclease mapping. The annealed RNA and probe were co-precipitated with ethanol, dried and then redissolved in water (18.5 μ l). To this solution were added buffer [0.1 M-Tris/HCl (pH 8.3)/0.15 M-KCl/10 mM- $MgCl_2$ /10 mM-dithiothreitol] and 1 mM each of the four non-radioactive deoxynucleoside triphosphates. The probe was extended with reverse transcriptase (16 units) for 30 min at 42°C. Then the solution was diluted to 100 μ l with water, and the products were precipitated with ethanol, washed with 95% (v/v) ethanol and dried. They were then analysed by electrophoresis on a denaturing polyacrylamide gel (6%), with the dideoxy DNA sequencing reaction products of the M13 clone as marker.

Table 1. Probes used for transcriptional mapping with S_1 nuclease and for primer extension analysis

Probe no.	Start*	Restriction-endonuclease* digestion site	Length† (bases)
1	3828	<i>Aha</i> II, 3489	387
2	3663	<i>Aha</i> II, 3489	222
3	3828	<i>Sma</i> I, 3371	505
4	3663	<i>Sma</i> I, 3371	340
5	3457	<i>Cla</i> I, 3282	223
6	3418	<i>Cla</i> I, 3282	183
7	3457	<i>Sma</i> I, 3371	134
8	3418	<i>Sma</i> I, 3371	95
9	8775	<i>Sma</i> I, 8363	460
10	8775	<i>Bgl</i> II, 8577	246
11	8163	<i>Hae</i> II, 7878	333
12	8232	<i>Hae</i> II, 7878	402

* Position in nucleotide sequence given by Falk *et al.* (1985).

† All probes include an additional 48 bases derived from the primer (17 bases) and M13 DNA adjacent to the *Sma*I-digestion site (31 bases) (Messing & Vieira, 1982).

Results and discussion

Location of 5'-end of *atp*-operon mRNA

The site of initiation of transcription of the *Rsp. rubrum atp* operon was determined by two independent methods, mapping with *S*₁ nuclease and

primer extension analysis. Initial *S*₁-nuclease mapping experiments were conducted with probes covering the region on the 5' side of the δ -subunit gene (probes 1-4; Table 1). Full length protection was obtained for these probes, showing that the site of transcriptional initiation lay outside this region. In order to map this initiation site probes 5

Fig. 2. Mapping with *S*₁ nuclease of the 5'-end of *Rsp. rubrum atp*-operon mRNA

In (a) and (b) tracks are designated as follows: track M, size (bases) markers (being an *Msp*I-endonuclease digest of plasmid pBR322, end-labelled with [³²P]dCTP in the presence of the Klenow fragment of DNA polymerase and non-radioactive GTP); track a, 'prime-cut' probe annealed to *Rsp. rubrum* RNA and then digested with *S*₁ nuclease; track b, as a, without RNA; track c, 'prime-cut' probe alone. (a) Probe 5 (see Table 1). Arrows indicate the size of the probe (track c), 223 bases, and the size of the protected fragment (track a), 109 bases (b) Probe 6 (see Table 1). Arrows indicate the intact probe (track c), 183 bases, and in track a the size of the protected fragment, 70 bases.

and 6 (Table 1), which extend even further to the 5' side of the gene for the δ -subunit, were employed. The results of this experiment, shown in Fig. 2, indicate a 5'-end around nucleotide 3348 (Falk *et al.*, 1985). It is notable that a single band was not obtained, but a series of bands was produced, each differing by one nucleotide in length from its neighbour. A similar phenomenon was also observed in S_1 -nuclease mapping of transcripts from *Rps. blastica* (Tybulewicz *et al.*, 1984; also see below).

In order to obtain independent confirmation of the position of this 5'-end, primer extension analyses were then performed with probes 7 and 8 (Table 1). They extend from nucleotides 3457 and 3418 respectively to the *Sma*I site at nucleotide 3371. They were annealed with *Rsp. rubrum* RNA, and after extension with reverse transcriptase gave a product extended by 24 bases from the *Sma*I site, as illustrated in the case of probe 7 in Fig. 3. This corresponds to a 5'-end for *Rsp. rubrum atp*-operon

mRNA at nucleotide 3348, thus supporting the result obtained by S_1 -nuclease mapping.

The sequence of the *Rsp. rubrum atp* operon has a non-coding segment of 140 nucleotides between the genes for the β - and ϵ -subunits. However, probes 11 and 12 (see Table 1), which encompass this region, gave only the expected full length protection in S_1 -nuclease mapping experiments and no other products. Thus the possibility that this region harbours a secondary promoter is eliminated.

In order to try and identify sequences that might serve as promoters in *Rsp. rubrum*, the DNA sequence upstream of this site was compared with the consensus of *E. coli* promoters. As shown in Fig. 4, this comparison showed the presence of a related sequence on the 5' side of transcriptional initiation base. However, in the absence of information concerning the binding of *Rsp. rubrum* RNA polymerase to this region, it is not known if these sequences represent a true promoter. Also, no

Fig. 3. Primer extension analyses to determine the *atp*-operon transcriptional initiation site. Track a, *Rsp. rubrum* RNA annealed with 'prime-cut' probe 7 (see Table 1) and extended with reverse transcriptase. Track b, 'prime-cut' probe 7 treated with reverse transcriptase in the absence of RNA. Track c, 'prime-cut' probe 7 only. The arrows indicate the sizes of the probe in track c (134 bases) and the extended probe in track a (158 bases). On the left of the diagram is shown the dideoxy sequence pattern of the region showing the *Sma*I-endonuclease site at which the probe ended (see Table 1).

Fig. 4. Comparison of *E. coli* promoter sequences with sequences on the 5' sides of the *Rsp. rubrum atp* operon and of an unidentified reading frame in *Rsp. rubrum*

The *E. coli* consensus sequence is taken from Pribnow (1978). The rightmost box is the site of initiation of transcription. This has been experimentally established for the *Rsp. rubrum atp*-operon cluster as described in the present paper and is nucleotide 3348. The numbers between boxes indicate the number of intervening nucleotides. For the sequences of *Rsp. rubrum* see Falk *et al.* (1985).

Fig. 5. Determination of the site of termination of transcription of the *atp*-operon mRNA by mapping with S_1 nuclease. In (a) and (b) tracks are labelled as follows: track M, size markers (being an *Msp*I-endonuclease digest of pBR322 DNA end-labelled with [32 P]dCTP (see Fig. 2); track a, 'prime-cut' probe annealed with *Rsp. rubrum* RNA and then digested with S_1 nuclease; track b, as a, without RNA; track c, 'prime-cut' probe alone. (a) Probe 10 (see Table 1). Arrows indicate the size of the probe in track c (246 bases) and the size of the protected fragment in track a (159 bases). (b) Probe 9 (Table 1). Here arrows again show intact probe in track c (460 bases) and the size of the protected fragment in track a (373 bases).

sequences could be identified that are related either to elements upstream of the *Rps. blastica atp* operon (Tybulewicz *et al.*, 1984) or to a non-coding region preceding the gene for *Rsp. rubrum* ribulose-phosphate carboxylase (Nargang *et al.*, 1984).

Studies of promoters in a wide range of other species, including *Bacillus subtilis* (Johnson *et al.*, 1983), *Klebsiella aerogenes* (Beynon *et al.*, 1983) and *Rhizobium meliloti* (Sundaresan *et al.*, 1983), have shown a large range of sequences involved in

Fig. 6. Potential secondary structure at the 3'-end of the *Rsp. rubrum* *atp* operon

(a) Secondary structure, including the transcriptional termination base (arrowed). (b) Other potential structures of unknown function between the termination codon of the gene for F_1 -ATPase ϵ -subunit and the transcriptional terminator.

promotion of transcription in other prokaryotes. This presumably reflects a divergence in the promoter-binding factor of RNA polymerases in these prokaryotes.

Termination of the *Rsp. rubrum* *atp* mRNA

Probes 9 and 10 (see Table 1) covering a region of DNA beyond the 3'-end of the gene for the ϵ -subunit of *Rsp. rubrum* F_1 -ATPase were used to map the 3'-end of the *atp*-operon mRNA. As shown in Fig. 5, the protected fragments obtained with these probes again had ragged ends, possibly due to 'breathing' at the end of the DNA-RNA duplex. The longest fragments correspond to a termination site at nucleotide 8736.

This nucleotide falls just before a run of thymidylate residues immediately adjacent to a region of dyad symmetry able to form a stem-loop structure as shown in Fig. 6. In *E. coli* these features are characteristic of *rho*-independent transcriptional terminators (Pribnow, 1978). The *atp* operon in *Rps. blasticus* also appears to terminate at a related structure (Tybulewicz *et al.*, 1984). Thus it appears that *rho*-independent transcriptional terminators are conserved in a wide range of species.

Between the termination codon of the gene for the ϵ -subunit and this potential structure are to be found two other regions with the potential to form stable stem-loop structures (Fig. 6). However, no transcriptional termination was evident in these regions, and their functions (if any) are unknown.

G. F. is supported by a grant from the Swedish Ministry of Education.

References

- Berk, A. J. & Sharp, A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1274-1278
- Beynon, J., Cannon, M., Buchanan-Wollaston, V. & Cannon, F. (1983) *Cell (Cambridge, Mass.)* **34**, 665-671
- Chory, J. & Kaplan, S. (1982) *J. Biol. Chem.* **257**, 15110-15121
- Falk, G., Hampe, A. & Walker, J. E. (1985) *Biochem. J.* **228**, 391-407
- Farrell, P. J., Deininger, P. L., Bankier, A. & Barrell, B.G. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1565-1569
- Gay, N. J. (1984) *J. Bacteriol.* **158**, 820-825
- Gay, N. J. & Walker, J. E. (1981) *Nucleic Acids Res.* **9**, 3919-3926
- Johnson, W. C., Moran, C. P. & Losick, R. (1983) *Nature (London)* **302**, 800-804
- Messing, J. & Vieira, J. (1982) *Gene* **19**, 269-276
- Nargang, F., McIntosh, L. & Somerville, C. (1984) *Mol. Gen. Genet.* **193**, 220-224
- Porter, A. C. G., Brusilow, W. S. A. & Simoni, R. D. (1983) *J. Bacteriol.* **155**, 1271-1278
- Pribnow, D. (1978) in *Biological Regulation and Development*, vol. 1 (Goldberger, R. F., ed.), pp. 219-277, Plenum Press, New York
- Sundaresan, V., Jones, J. D. G., Ow, D. W. & Ausubel, F. M. (1983) *Nature (London)* **301**, 728-732
- Tybulewicz, V. L. J., Falk, G. & Walker, J. E. (1984) *J. Mol. Biol.* **179**, 185-214
- Walker, J. E., Gay, N. J., Saraste, M. & Eberle, A. N. (1984a) *Biochem. J.* **224**, 799-815
- Walker, J. E., Saraste, M. & Gay, N. J. (1984b) *Biochim. Biophys. Acta* **768**, 164-200